Hájek Martin

Vysoce dostupný firewall

Iptables

http://www.netfilter.org/

Iptables je linuxový firewall, který umí zajistit několik věcí: filter umožňuje filtrovat síťový provoz na úrovni paketů pomocí pravidel (odkud, kam, kudy… paket cestuje), tabulka mangle umožňuje změnu nastavení paketů. Obě tyto tabulky nepředstavují problém pro spuštění na záložním počítači při pádu primárního. Stačí mít skript, který tato pravidla nastaví. Nejlepší by bylo, kdyby byl skript uložen na sdíleném síťovém disku, aby se při změně skriptu nemuselo upravovat více souborů.

Poslední je tabulka nat. Ta je problematická na zajištění vysoké dostupnosti. Nastavení pravidel může také sice proběhnout pomocí skriptu, ale při použití "maškarády", nebo přesměrování portů si firewall musí uložit informace o aktuálních spojení mezi počítačem ve vnitřní síti a serverem v internetu. Při pádu počítače se tato tabulka ztratí a všechna spojení se přeruší. Na záložním počítači je třeba nastavit pravidla v tabulce a navázat nová spojení. Nebo je nutné tyto tabulky synchronizovat již při běhu na obou dvou počítačích. Což není jednoduché.

V rozhovoru s jedním z tvůrců Iptables jsem se dočetl, že plánované Iptables2 by v sobě mělo údajně obsahovat zajištění vysoké dostupnosti, ale zatím není jasné na jak vysoké úrovni. (tato informace je tři roky stará ☹ novější informace jsem nezjistil)

Nalezená vysoce dostupná řešení:

- HeartBeat
- Keepalive
- Piranha a Cluster manager
- UltraMonkey
- DRBD - (virtuální síťový disk)

Heartbeat

http://www.linux-ha.org/HeartbeatProgram

Heartbeat je základní, nejjednodušší program na zajištění vysoké dostupnosti.

Funkce: Dva počítače mají nainstalovaný Heartbeat (primární stroj a sekundární záložní) Po nastartování systémů na primárním počítači Heartbeat nastaví druhou(může jich být i více) IP adresu na síťovém rozhraní, pod touto IP adresou se spustí požadované služby (na tuto adresu se směrují požadavky klientů…).

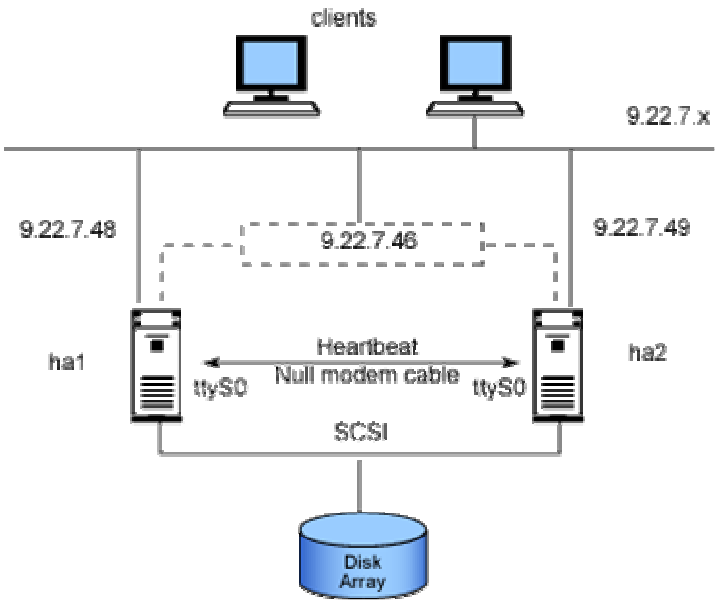
Sekundární uzel posílá po nastavené době dotazy primárnímu uzlu, zda je naživu (dotazy přes síť, nebo přes COM port) pokud se primární uzel neozve do nastavené doby vydá varování a po chvíli ho prohlásí za mrtvého. Pak sekundární uzel přebere jeho roli. Stejně jako výše: Přidání druhé IP adresy na síťové rozhraní a spuštění služeb. Doporučená doba po které je uzel prohlášený za mrtvého je zhruba 20sec. Více je zbytečné a méně může vést ke zbytečnému přepínání uzlů (to, že uzel neodpovídá může být způsobeno např. krátkodobým vytížením systému…)

Je možné nastavit, aby se po opětovném spuštění primárního uzlu služby na sekundárním uzlu ukončily a spustily na opět na primárním. Nebo je toto rozhodnutí možné nechat na správci sítě. (např. počkat na nižší provoz sítě...)

Obrázek ukazuje jedno z možných zapojení Heartbeatu. Jsou zde dva počítače ha1 a ha2. Každý má svojí IP adresu 9.22.1.48 (49). Aktivní uzel na kterém běží služby dostane ještě druhou IP adresu 9.22.7.46 na kterou jsou směrovány požadavky klientů. Zjišťování, zda

je uzel naživu se zde děje pomocí COM1 portu. Sdílení dat je přes externí společný disk spojený SCSI sběrnicí.

HB poskytuje vysoce dostupné řešení. Je přenosný na každou distribuci linuxu. Je jednoduchý na instalaci a nastavení.



Nevýhoda toho řešení je, že při pádu uzlu se služby spouští znovu (ztráta spojení, neuložených dat…), čeká se poměrně dlouhou dobu než je uzel prohlášený za mrtvého. během této doby jsou pakety, které jsou k němu směrovány, zahozeny. Heartbeat je navržený jen pro dva uzly. HB samotný neumí přímo sdílet pro oba uzly stejná data. K tomu slouží další programy jako např. DRBD. Nebo hardwarové společné datové úložiště (viz níže).

Pro částečné zajištění vysoké dostupnosti Iptables je možné si napsat skript na nastavení pravidel a ten při pádu primárního uzlu spustit na záložním.

HB jsem nainstaloval, ale testoval jsem ho jen na jednom počítači ☹. Vyzkoušel jsem, jeho základní nastavení, jak přidá IP adresu na síťové rozhraní a na ní spustí služby (zkoušel jsem apache a malý skript na nastavení Iptables)

HB se hodí jako základní řešení pro zajištění vysoké dostupnosti na místech kde krátkodobý výpadek služeb (do 30sec) nic neohrozí.

Keepalived

http://keepalived.sourceforge.net

Je program podobný Heartbeatu, ale má několik rozšíření. Keepalived zjišťuje stav, dostupnost primárního uzlu pomocí několika běžných protokolů (např. http, nebo i pomocí nějakého jiného démona). Keepalive používá protokol VRRP, který zajistí kompletní spuštění služeb na záložním uzlu při pádu primárního. (Předání VIP…)

Výhoda oproti Heartbeatu je možnost použití více záložních uzlů s různými prioritami na přebírání služeb. Keepalived sám nezajišťuje sdílení dat. Opět DRBD, nebo hardwarové řešení…

Piranha a cluster manager

http://sourceware.org/piranha/

Piranha slouží k rozdělování požadavků klientů mezi více počítačů.
http://www.redhat.com/docs/manuals/enterprise/RHEL-AS-2.1-Manual/cluster-manager/
Cluster manager slouží z zajištění vysoké dostupnosti. Obě aplikace mohou pracovat současně a to přináší výhody z obou systémů.

Cluster manager: je to profesionální řešení. Podporuje použití až 8 uzlů, podporuje použití NFS(network file systém) a CIFS(common internet FS), má plně sdílený úložný systém, poskytuje plnou garanci sdílených dat, testuje nejen dostupnost celého serveru, ale i jednotlivých služeb (pro případ chyby v SW). Sdílení dat je pomocí externího sdíleného

disku: externí RAID pole (nejlépe se zdvojeným řadičem, připojené přes paralelní SCSI nebo Fibre chanel).

UltraMonkey

http://www.ultramonkey.org/

Ultra monkey nepřináší nic nového. Jen zastřešuje stávající řešení jako Heartbeat, LVS (systém na rozdělování zátěže) a Ldirector (monitorování a administrace LVS).

DRBD sw sdílený disk

http://www.drbd.org/

DRBD je patch do jádra. Jde o něco podobného jako síťový RAID-1.

Měl by být přenosný na jakoukoliv distribuci linuxu.

Instalace a konfigurace DRBD nevypadá složitě. Stačí upravit pár řádek jako IP adresy serverů, jaký disk se má použít na sdílení… Když je DRBD zprovozněné samo se synchronizuje na obou dvou počítačích, objeví se nové zařízení /dev/drbd0 které je dostupné z obou uzlů. Je zde zajištěná konzistence dat (aby dva uzly nezapisovaly současně, nečetly stará data… stejné problémy jako např. u databází)

DRBD jsem nezkoušel, protože na internetu jsem si přečetl, že si tím člověk může přemazat data na disku (i s tou jednoduchou konfigurací ☺) a to jsem nechtěl riskovat.

Hardwarové sdílení dat

Další možností jak sdílet data je hardwarové řešení jako je externí Raid pole se dvěma řadiči a propojení k počítačům zajišťuje paralelní SCSI sběrnice nebo Fibre chanel.

Závěr

Jak nutné je zajistit absolutní vysokou dostupnost? Jak často dochází k chybám na serveru? Spojení se může ztratit i "někde v internetu"… Nestačí se spokojit se ztrátou aktuálních spojení a nedostupností po dobu cca 30 sekund? Pro server běžné firmy stačí použít jedno z výše uváděných řešení (i to nejjednodušší heartbeat) Pokud nějaká společnost např. poskytovatel internetu bude potřebovat absolutní vysokou dostupnost použije nějaké placené řešení, které bude obsahovat prověřenou kombinaci hardwaru a softwaru. Tento software nebude nejspíš iptables (může z něho vycházet), ale něco přizpůsobené na desítky tisíc navázaných spojení v jednom okamžiku.